CDDIS Global Data Center Technical Report 2013

C. Noll

NASA Goddard Space Flight Center, Code 690.1
Greenbelt, MD 20771 USA
Carey.Noll@nasa.gov

1 Introduction

The Crustal Dynamics Data Information System (CDDIS) is NASA's data archive and information service supporting the international space geodesy community. For over 30 years, the CDDIS has provided continuous, long term, public access to the data (mainly GNSS-Global Navigation Satellite System, SLR-Satellite Laser Ranging, VLBI-Very Long Baseline Interferometry, and DORIS-Doppler Orbitography and Radiopositioning Integrated by Satellite) and products derived from these data required for a variety of science observations, including the determination of a global terrestrial reference frame and geodetic studies in plate tectonics, earthquake displacements, volcano monitoring, Earth orientation, and atmospheric angular momentum, among others. The specialized nature of the CDDIS lends itself well to enhancement to accommodate diverse data sets and user requirements. The CDDIS is one of NASA's Earth Observing System Data and Information System (EOSDIS) distributed data archive centers; EOSDIS data centers serve a diverse user community and are tasked to provide facilities to search and access science data and products.

The CDDIS serves as one of the primary data centers and core components for the geometric services established under the International Association of Geodesy (IAG), an organization that promotes scientific cooperation and research in geodesy on a global scale. The system has supported the International GNSS Service (IGS) as a global data center since 1992. The CDDIS activities within the IGS during 2013 are summarized below; this report also includes any recent changes or enhancements made to the CDDIS.

2 System Description

The CDDIS archive of IGS data and products are accessible worldwide through anonymous ftp. The CDDIS is located at NASA's Goddard Space Flight Center (GSFC) and is available to users 24 hours per day, seven days per week.

The CDDIS computer system is fully redundant with the primary and secondary/failover system. Each system utilizes a distributed functionality (incoming, outgoing, processing, database, and map servers) and is configured with a local backup system as well as a full backup system located in a third building at GSFC. The archive is equipped with a multi-Tbyte RAID storage system and is scaled to accommodate future growth. All ftp and web access is performed on the outgoing server. Data centers, stations, and analysis centers push files to the CDDIS incoming server. Processing of incoming files for the on-line archive is performed in a separate environment that also includes a database server for managing metadata extracted from incoming data.

3 Archive Content

As a global data center for the IGS, the CDDIS is responsible for archiving and providing access to GNSS data from the global IGS network as well as the products derived from the analyses of these data in support of both operational and working group/pilot project activities. The CDDIS archive is approximately 7.9 Tbytes in size of which 7.5 Tbytes are devoted to GNSS data, products (300 Gbytes),

and ancillary information. All data and products are accessible through subdirectories of *ftp://cddis.gsfc.nasa.gov/gnss* (a symbolic link to *ftp://cddis.gsfc.nasa.gov/gps*).

3.1   GNSS Tracking Data

The user community has access to GNSS data available through the on-line global data center archives of the IGS. Over 50 operational and regional IGS data centers and station operators make data (observation, navigation, and meteorological) available in RINEX format to the CDDIS from selected receivers on a daily, hourly, and sub-hourly basis. The CDDIS also accesses the archives of the other three IGS global data centers, Scripps Institution of Oceanography (SIO) in California, the Institut Géographique National (IGN) in France, and the Korea Astronomy and Space Science Institute (KASI) to retrieve (or receive) data holdings not routinely transmitted to the CDDIS by an operational or regional data center. Table 1 below summarizes the types of IGS operational GNSS data sets archived at the CDDIS.

Table 1a: GNSS Data Type Summary.

Data Type	Sample Rate	Data Format	Available On-line
Daily GNSS	30 sec.	RINEX and compact RINEX	Since 1992
Hourly GNSS	30 sec.	Compact RINEX	10+ years
High-rate GNSS	1 sec.	Compact RINEX	Since May 2001
Satellite GPS	10 sec.	Compact RINEX	Since 2002

Table 1b: GNSS Data Archive Summary for 2013.

Data Type	Avg. No. Sites/Day	Avg. Volume/Day	Total Volume/Year	Directory Location	Latency of Majority of Data
Daily GNSS	475	1100 Mb	400 Gb	*/gnss/data/daily*	Within 1 hour
Hourly GNSS	290	383 Mb	140 Gb	*/gnss/data/hourly*	Within 10 minutes
High-rate GNSS	155	1900 Mb	695 Gb	*/gnss/data/highrate*	Within 10 minutes
On-board GPS	1	0.5 Mb	200 Mb	*/gnss/data/satellite*	Within 10 days

Data, in RINEX V2.10 or V2.11 format, from GPS and GPS+GLONASS receivers are archived within the main GNSS directory structure */gnss/data*.

The CDDIS archives four major types/formats of GNSS data, all in RINEX format, as described in Table 1a. Daily RINEX data are quality-checked, summarized, and archived to public disk areas in subdirectories by year, day, and file type; the summary and inventory information are also loaded into an on-line database. Nearly 153K daily station days from 520 distinct GNSS receivers were archived at the CDDIS during 2013. A complete list of daily, hourly, and high-rate sites archived in the CDDIS can be found in the yearly summary reports at URL *ftp://cddis.gsfc.nasa.gov/reports/gnss/*.

Within minutes of receipt, the hourly GNSS files are archived to subdirectories by year, day, and hour. These data are retained on-line indefinitely; the daily files delivered at the end of the UTC day contain all data from these hourly files and thus can be used in lieu of the individual hourly files. A total of 360 hourly sites (4.5M files) were archived during 2013.

High-rate (typically 1-second sampling) GNSS data are archived in files containing fifteen minutes of data and in subdirectories by year, day, file type, and hour. Many of these data files are created from real-time streams. Data from 175 high-rate sites (4.8M files) were also archived in the CDDIS in 2013.

The CDDIS generates global broadcast ephemeris files (for both GPS and GLONASS) on an hourly basis. These files are derived from the site-specific ephemeris data files for each day/hour. These files are appended to a single file that contains the orbit information for all GPS and GLONASS satellites for the day up through that hour. The merged ephemeris data files are then copied to the day's subdirectory within the hourly data file system. Within 1-2 hours after the end of the UTC day, after sufficient station-specific navigation files have been submitted, this concatenation procedure is repeated to create the daily broadcast ephemeris files (both GPS and GLONASS), using daily site-specific navigation files as input. The daily files are copied to the corresponding subdirectory under the daily file system. Users can thus

download this single, daily (or hourly) file to obtain the unique navigation messages rather than downloading multiple broadcast ephemeris files from the individual stations.

The CDDIS generates and updates status files, (*/gnss/data/daily/YYYY/DDD/YYDD.status*) that summarize the holdings of daily GNSS data. These files include a list of stations. The archive status files of CDDIS GNSS data holdings reflect timeliness of the data delivered as well as statistics on number of data points, cycle slips, and multipath. The user community can receive a snapshot of data availability and quality by viewing the contents of such a summary file.

The CDDIS successfully submitted a proposal to the IGS Multi-GNSS Experiment (MGEX) call for proposals for archive and distribution of data and products. During 2013 the CDDIS continued the archiving of data from participating multi-GNSS receivers as well as products derived from the analysis of these data. The data include newly available signals (e.g., Galileo, QZS, SBAS, and Beidou). The summary of the MGEX data holdings at the CDDIS is shown in Table 2 below. Daily status files are also provided that summarize the MGEX data holdings; however, data quality information available for operational GNSS data holdings is not available through this software.

Table 2: GNSS MGEX Data Archive Summary for 2013.

Data Type	Avg. No. Sites/Day	Avg. Volume/Day	Directory Location
Daily GNSS	100	450 Mb	*/gnss/data/campaign/mgex/daily*
Hourly GNSS	50	90 Mb	*/gnss/campaign/mgex /data/hourly*
High-rate GNSS	35	850 Mb	*/gnss/campaign/mgex /data/highrate*

The GPS Directorate conducted the first test transmission of CNAV (civilian navigation) messages on L2C- and L5-capable satellites in June 2013. The civilian navigation message types will provide improved navigation through the use of the new L2C and L5 civilian frequencies. The German Aerospace Center (DLR) and the University of New Brunswick (UNB) teamed up to collect the CNAV data with five multi-GNSS receivers in the U.S., Canada, South Africa, Singapore, and Australia. The team at DLR and UNB coordinated archive of the data with the CDDIS in support of the IGS MGEX.

Data from the permanent network of 20 GPS monitoring stations maintained by the National Geospatial-Intelligence Agency (NGA) were archived in a special area of the CDDIS for review by select IGS analysis centers. The data span the 2010-2012 time frame. Once review has been completed the data will be archived in the operational directories.

3.2   IGS Products

The CDDIS routinely archives IGS operational products (daily, rapid, and ultra-rapid orbits and clocks, and weekly ERP and station positions) as well as products generated by IGS working groups and pilot projects (ionosphere, troposphere, real-time clocks). Table 3 below summarizes the GNSS products available through the CDDIS. The CDDIS currently provides on-line access through anonymous ftp to all IGS products generated since the start of the IGS Test Campaign in June 1992 in the file system */gnss/products*; products from GPS+GLONASS products are available through this filesystem. Products derived from GLONASS data only continued to be archived at the CDDIS in a directory structure within the file system */glonass/products*.

Table 3: GNSS Product Summary.

Product Type	Number of ACs/AACs	Volume	Directory
Orbits, clocks, ERP, positions	14+Combinations	1.2 Gb/week	*/gnss/products/WWWW (GPS, GPS+GLONASS)* */glonass/products/WWWW (GLONASS only)*
Troposphere	Combination	2.6 Mb/day, 940 Mb/year	*/gnss/products/troposphere/YYYY*
Ionosphere	4+Combination	4 Mb/day, 1.5 Gb/year	*/gnss/products/ionex/YYYY*
Real-time clocks	Combination	6.0 Mb/week	*/gnss/products/rtpp/YYYY*

Note: *WWWW*=4-digit GPS week number; *YYYY*=4-digit year

The CDDIS also continues to archive combined troposphere estimates in directories by GPS week. Global ionosphere maps of total electron content (TEC) from the IONEX AACs are archived in subdirectories by year and day of year. Real-time clock comparison products have been archived at the CDDIS in support of the IGS Real-Time Pilot Project since 2009.

3.3 Supporting Information

Daily status files of GNSS data holdings, reflecting timeliness of the data delivered as well as statistics on number of data points, cycle slips, and multipath, continue to be generated by the CDDIS. By accessing these files, the user community can receive a quick look at a day's data availability and quality by viewing a single file. The daily status files are available through the web at URL *ftp://cddis.gsfc.nasa.gov/reports/gnss/status*. The daily status files are also archived in the daily GNSS data directories.

In preparation for the analysis center's second reprocessing campaign, the CDDIS has developed site-specific reports detailing missing data. Station operators and operational data centers can consult these lists (*ftp://cddis.gsfc.nasa.gov/gnss/data/daily/reports/missing*) and if available, supply missing files to the CDDIS for inclusion in the global data center archives.

Ancillary information to aid in the use of GNSS data and products are also accessible through the CDDIS. Weekly and yearly summaries of IGS tracking data (daily, hourly, and high-rate) archived at the CDDIS are generated on a routine basis. These summaries are accessible through the web at URL *ftp://cddis.gsfc.nasa.gov/reports/gnss*. The CDDIS also maintains an archive of and indices to IGS Mail, Report, Station, and other IGS-related messages.

4 System Usage

Figure 1 summarizes the usage of the CDDIS for the retrieval of GNSS data and products in 2013. This figure illustrates the number and volume of GNSS files retrieved by the user community during 2013, categorized by type (daily, hourly, high-rate, MGEX data, products). Nearly 370 million files (over 50 Tbytes), excluding robot downloads, were transferred in 2013, with an average of nearly 30 million files per month. Figure 2 illustrates the profile of users accessing the CDDIS IGS archive during 2013. The majority of CDDIS users are from hosts in Europe, Asia, and North America.
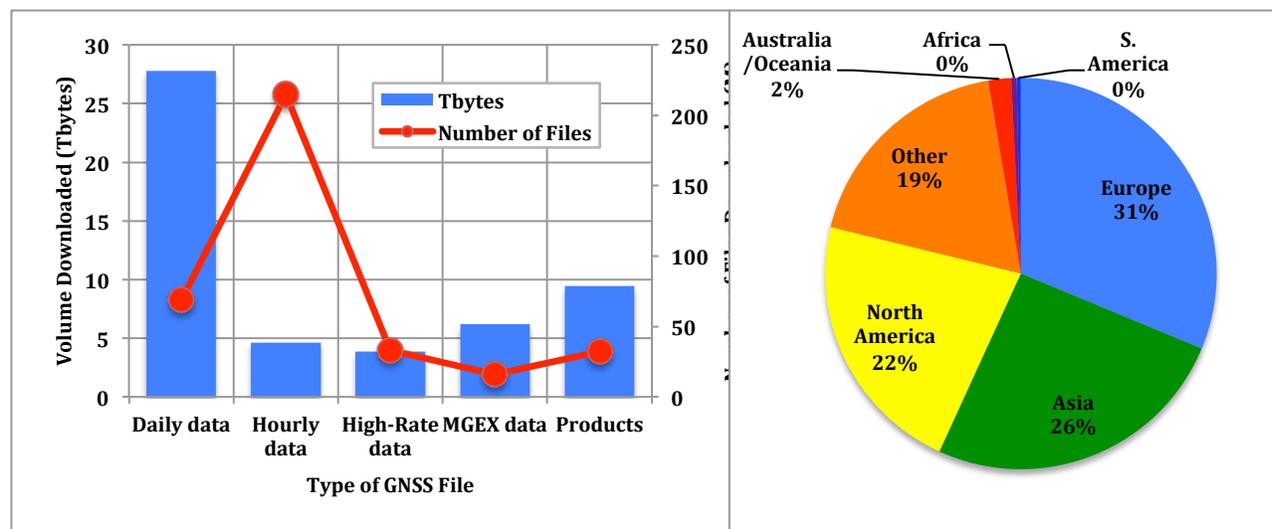

Figure 1: Number and volume of GNSS files transferred from the CDDIS in 2013.

Figure 2: Geographic distribution of IGS users of the CDDIS in 2013.

5   Recent Developments

In support of the IGS Real-Time Pilot Project (RTPP) and Real-Time IGS Service, the CDDIS has configured a server to provide a real-time streaming capability at GSFC. The CDDIS installed the Ntrip (Networked Transport of RTCM via Internet Protocol) software and successfully tested obtaining product streams from BKG and providing access to streams to authorized users. A module was developed by EOSDIS colleagues to plug into Ntrip that interfaces with an established User Registration System (URS) at NASA GSFC. The module was specifically developed to easily interface with multiple user verification systems and was given back to the community for possible future use. Testing continues on the CDDIS caster installation and the Ntrip user registration module while the CDDIS role in the Real-Time IGS Service is established.

The CDDIS has recently made modifications to the metadata extracted from incoming data and product files pushed to its archive. These enhancements have facilitated cross discipline data discovery by providing information about CDDIS archive holdings to other data portals such as Earth Observing System (EOS) Clearinghouse (ECHO) and future integration into the Global Geodetic Observing System (GGOS) portal. The staff has begun a metadata evolution effort, re-designing the metadata extracted from incoming data and adding information that will better support EOSDIS applications such as ECHO and the metrics collection effort.

The CDDIS began discussions with the International Council for Science (ICSU) World Data System (WDS) in 2011 to apply for membership. The CDDIS was previously accepted through the ILRS application as a network member. The WDS strives to enable open and long-term access to multidisciplinary scientific data, data services, products and information. The WDS works to ensure long-term stewardship of data and data services to a global scientific user community. In March 2013, the ICSU WDS Scientific Committee accepted the CDDIS application for regular membership. A Letter of Agreement between the CDDIS and ICSU was drafted and signed by C. Noll (on behalf of the CDDIS) and D. Lowe (on behalf of EOSDIS, CDDIS funding organization).

6   Publications

The CDDIS staff attended several conferences during 2013 and presented papers on or conducted demos of their activities within the IGS, including:

> C. Noll, M. Dube, N. Pollack, L. Tyahla, P. Michael. "Improvements in Space Geodesy Data Discovery at the CDDIS", Abstract IN31C-1520 presented at 2013 Fall Meeting, AGU, San Francisco, Calif., 09-13 Dec.

Electronic versions of this poster and other publications can be accessed through the CDDIS on-line documentation page on the web at URL *http://cddis.gsfc.nasa.gov/reports.html*.

7   Future Plans

The CDDIS will continue to support the IGS MGEX. The experiment is an excellent opportunity to prepare the data centers for archive of data in RINEX V3. The CDDIS will coordinate with the Data Center Working Group and other IGS data centers to introduce RINEX V3 data into the "operational" GNSS data directory structure, making it easier for users to access these data.

The CDDIS plans to make its real-time caster operational in the coming year as part of the IGS Real-Time Service. Possible future activities in the real-time area include capturing the streams for generation of 15-minute high-rate files for archive at the CDDIS.

In 2013, the IGS analysis centers will begin providing products for the second IGS reprocessing campaign (repro2). The CDDIS will provide support through upload of files from the ACs and online archive of these products (*/gnss/products/WWWW/repro2*).

Work on an update of the CDDIS website began in 2013. In addition to a refresh of the appearance of the website, the content was reviewed and updated. Applications that are under development for data discovery (e.g., a CDDIS implementation of GSAC and a site log browser) will be completed in early 2014 and integrated into the new CDDIS website. The update is planned for completion in early 2014.

The CDDIS has cooperated in the development of Geodetic Seamless Archive Centers (GSAC) with colleagues at UNAVCO, SIO, and the University of Nevada at Reno. The activity provides web services to facilitate data discovery within and across participating archives. The prototype implementation of these GSAC web services at the CDDIS was developed; plans are to incorporate the application in the revised CDDIS website in early 2014.

Funding has been identified in 2014 to procure a system refresh for the CDDIS. The CDDIS system engineer will review current and near-term requirements and develop a hardware procurement strategy.

8   Contact Information

To obtain more information about the CDDIS IGS archive of data and products, contact:

Ms. Carey E. Noll	Phone:	(301) 614-6542
Manager, CDDIS	Fax:	(301) 614-6015
Code 690.1	E-mail:	*Carey.Noll@nasa.gov*
NASA GSFC	WWW:	*http://cddis.gsfc.nasa.gov*
Greenbelt, MD 20771		*http://cddis.nasa.gov*

9   Acknowledgments

References

C. Noll, The Crustal Dynamics Data Information System: A resource to support scientific analysis using space geodesy, Advances in Space Research, Volume 45, Issue 12, 15 June 2010, Pages 1421-1440, ISSN 0273-1177, DOI: 10.1016/j.asr.2010.01.018.

C. Noll, Y. Bock, H. Habrich and A. Moore, "Development of data infrastructure to support scientific analysis for the International GNSS Service", Journal of Geodesy, Feb 2009, pages 309-325, DOI 10.1007/s00190-008-0245-6.